<u>Summary File 3: Data Note 4 – Updated December 2002</u>

In July 2002, the Census Bureau issued the following Data Note 4 regarding the Census 2000 Summary File 3 (SF3) data:

> The Census Bureau is aware there may be a problem or problems in the employment-status data of Census 2000 Summary File 3 (including tables P38, P43-46, PCT35, P149A-1, P150A-I, PCT35, PCT69A-1, and PCT 70A-1). The labor force data for some places where colleges are located appear to overstate the number in the labor force, the number unemployed, and the percent unemployed, probably because of reporting or processing errors. The exact cause is unknown, but the Census Bureau will continue to research the problem.

Our further research into this "college-town" issue indicates that the problem extended beyond places with colleges to the country in general. We learned that it stems from the tendency of many working-age people living in civilian noninstitutional group quarters (GQ), such as college dormitories, worker dormitories, and group homes (for the mentally ill or physically handicapped), to exhibit a particular pattern of entries to the employment questions in Census 2000.[1] We now estimate that the pattern affected the employment data for about 15 percent of the civilian noninstitutional GQ population 16 years of age and over in the United States, or around 500,000 people. It had an impact on the Census 2000 labor force statistics for the entire country, but its effects were most visible and substantial for places, such as college towns, with high concentrations of people living in civilian noninstitutional group quarters.

In Census 2000, the majority of people in the GQ population were enumerated by the Individual Census Report (ICR) form, which collected employment data in a battery of six questions (questions 23, 27a-e). The responses to these questions were captured and fed into a set of rules (called the Employment Status Recode (ESR) edit) that used the combined information from all six questions to assign each person to one of the following four employment-status categories: not in universe (all people less than 16 years old), employed, unemployed, and not in labor force.

For a significant segment of the GQ population, a so-called "3/3" response pattern was entered into the ESR edit.[2] This pattern is shown in the following table:

---

[1] The pattern also appeared frequently for people in institutional group quarters, such as prisons and juvenile institutions, but, because of the way employment categories are defined, it had no impact on the employment data for these people.

[2] "3/3" refers to the fact that the responses to the first three questions, which appeared on page 4 of the ICR, are all missing; and those responses to the last three questions, which were on page 5 of the ICR, are all "yes."

3/3 Input Pattern from ICR forms

Question Number on ICR	Question Wording	Entry
23	LAST WEEK, did you do ANY work for either pay or profit?	Missing
27a	LAST WEEK, were you on layoff from a job?	Missing
27b	LAST WEEK, were you TEMPORARILY absent from a job or business?	Missing
27c	(For people on layoff) Have you been informed that you will be recalled to work within the next 6 months OR been given a date to return to work?	Yes
27d	Have you been looking for work during the last four weeks?	Yes
27e	LAST WEEK, could you have started a job if offered one, or returned to work if recalled?	Yes

The 3/3 pattern represents an incomplete set of information, since entries to the first three questions are missing. The ESR edit assigned people with this pattern to the "unemployed" category, because the edit had three built-in assumptions:
> 1) the respondents saw and reacted to each and every question in the employment series;
> 2) the 3/3 pattern represented the faithful recording of actual responses (or non-responses) to the questions; and
> 3) people who responded in this manner were more likely to meet the official criteria for the "unemployed" category than for any other category.[3]

Our research has revealed that most of the GQ cases with the 3/3 pattern may not have met one of the first two assumptions. We are still investigating, but we think that, in most cases, the pattern resulted from anomalies in the data collection or processing systems. Unfortunately, we cannot test our hypothesis by comparing the 3/3 pattern with actual reports from the respondents. The images of the filled-out ICR's will not be accessible until the completion, in 2006 at the earliest, of the Census Bureau's project to image the forms for delivery to the National Archives.

The potential effect of the ESR outcome for the 3/3 pattern is to increase the count of unemployed people at the expense of the counts of the employed and the not-in-labor-force

---

[3] They reported that they were looking for work and could have started a job last week. Because they did not report whether they had a job last week (people with a job are classified as "employed"), it is reasonable to classify them as "unemployed."

groups. We have done some research to estimate the potential impact of the phenomenon on the labor force data for the nation as a whole. Our preliminary estimates are that it may have incorrectly decreased the number of employed people by about 235,000 (the number of employed in SF3 was 129.7 million), reduced the number of people not in the labor force by 285,000 (SF3 figure of 78.3 million), increased the number of unemployed by 519,000 (SF3 figure of 7.9 million), and raised the unemployment rate by 0.4 percentage point (SF3 figure was 5.8 percent).

Comparatively, the impact of the phenomenon on areas below the national level may be much greater, depending upon the relative size of the GQ population within the given area. The Census 2000 unemployment rate for the city of Williamsburg, Virginia, for example, was 41.7 percent. Our research indicated that this rate resulted primarily from the prevalence of the 3/3 pattern among residents of college dormitories, who make up a large percentage of the city's population.

We will continue our research and report on further findings as they become available.